

Kathrin Steyer

Kollokationen in deutschen Wörterbüchern und in der deutschen Wörterbuchforschung

1	Vorbemerkung	5	Erklärungsansätze und Visionen.
2	Kollokationen in der deutschsprachigen (Wörterbuch-)Forschung		Kookkurrenzen als Basis für Kollokationswörterbücher
3	Neues Kollokationswissen. Korpora, Ressourcen, Portale	6	Nachbemerkung
4	Korpuslinguistische Fundierung? Kollokationen in deutschen Wörterbüchern	7	Bibliografie
		7.1	Wörterbücher
		7.2	Internetadressen
		7.3	Sekundärliteratur

Abstract

This contribution article focuses on German-language collocation research and lexicographic practice from a corpus linguistic perspective. Although there is no dictionary called "Deutsches Kollokationswörterbuch" (German collocation dictionary), the collocation perspective acquires increasing popularity in linguistic research and dictionary work in the German-speaking area. On the one hand, this tendency is due to the growing number of studies dealing with German as a contrast language and works on foreign language didactics. On the other hand, powerful electronic resources such as large corpora and lexical databases, which are nowadays available for the German language, are recognised as a valuable empirical basis. Nevertheless, the application of novel corpus linguistic methods in lexicographic practice is still unsatisfactory. Therefore, this article concludes by discussing innovative aspects of corpus linguistic empirical research on the basis of collocations. These ideas are presented as an incentive for further research as well as practical application.

1 Vorbemerkung

Es gibt derzeit noch kein lexikografisches Produkt mit dem Namen „Deutsches Kollokationswörterbuch“. Das ist zunächst festzuhalten und möglicherweise zu bedauern. Worüber wird also in diesem Beitrag zu reden sein? Es wird zu reden sein über die zahlreichen, aber immer noch zu wenig exponierten Kollokationsangaben in deutschen Wörterbüchern, über den Nutzen von korpusanalytischen Verfahren, über die noch zu leistende korpuslinguistische Grundlagenforschung und über den schwierigen Weg all dieser Befunde und Ergebnisse in die lexikografischen Werkstätten. Die heutigen Möglichkeiten der Analyse sprachlicher Massendaten stellen eine völlig neue Herausforderung sowohl für die linguistische Interpretation der Daten als auch für die lexikografische Bearbeitung dar. Und es hat sich gezeigt, dass der Transfer von korpusanalytischen Erkenntnissen in anwendungsnahe Bereiche wie Lexikografie und Fremdsprachendidaktik über mehr Vermittlungsstufen erfolgen muss, als wir dies in der Anfangseuphorie über die Entdeckung des neuen Planeten Korpuslinguistik wahrhaben wollten.

Die im Beitrag angeführten Beispiele und Analysen basieren auf Daten, die mit Hilfe der am IDS entwickelten Kookkurrenzanalyse in allen IDS-Korpora der geschriebenen Sprache ermittelt wurden. Das vor allem dem 5. Kapitel zu Grunde liegende Beispiellemma *Grund* ist mit einem Vorkommen aller Wortformen von über 600 000 Treffern dem hochfrequenten Bereich zuzurechnen. Diese quantitative Dimension erlaubt auch seriöse Verallgemeinerungen auf den Usus.

2 Kollokationen in der deutschsprachigen (Wörterbuch-)Forschung

Die Kollokationsforschung im deutschsprachigen Raum ist relativ jung. Dies liegt nicht zuletzt an der „betriebsblinden Unschuld“ vieler muttersprachlich ausgerichteter Linguisten in der Vergangenheit (Hausmann 2004, 322).¹ Kollokationen galten – wenn sie denn überhaupt wahrgenommen wurden – in der Regel als banal oder trivial und auf Grund ihrer mangelnden Idiosynkrasie als uninteressant für linguistische Untersuchungen. Auch in aktuellen Einführungen zur germanistischen Linguistik sucht man den Terminus ‚Kollokation‘ oft noch vergeblich. Allenfalls finden sich Verweise auf Phraseologismen oder syntagmatische Strukturen (vgl. z.B. Busch/Stenschke 2007, Meibauer et al. 2007). Aus nichtmuttersprachlicher Sicht wurde die Kollokationsproblematik naturgemäß anders wahrgenommen, wie folgende zwei Verweise illustrieren sollen: „Wortschatzlernen ist Kollokationslernen“ titelt Hausmann einen seiner ersten programmatischen Aufsätze zu diesem Thema (1984, 395). Zöfgen bezeichnet Kollokationslücken als ein genuines Merkmal von Lernersprachen (2001, 271). Und so ist es nicht verwunderlich, dass die meisten der in den letzten Jahren erschienenen deutschsprachigen Monografien zum Thema ‚Kollokationen‘ einen bilingualen bzw. kontrastiven Hintergrund, zumeist mit fremdsprachendidaktischer Ausrichtung, haben.²

Forschungsansätze zur *deutschen Syntagmatik* können als einige der wenigen frühen Vorläufer deutschsprachiger Kollokationsforschung angesehen werden: In der Literatur werden vor allem Porzigs ‚wesenhafte Bedeutungsbeziehungen‘ (1934) und Coserius ‚lexikalische Solidaritäten‘ (1967) als solche einschlägigen Konzepte angesehen. Relevante Beiträge sowohl zum Status von Kollokationen im Lexikon (‚semantische Vereinbarkeit‘) als auch zu ihrer lexikografischen Beschreibung wurden am Zentralinstitut für Sprachwissenschaft in Berlin erarbeitet (vgl. Agricola 1982, Viehweger et al. 1977, Viehweger 1989). In diesem Umfeld sind ebenso innovative Wörterbücher entstanden. Das „Wörterbuch der

¹ Mit dem Erscheinen mehrerer an den Bedürfnissen von Fremdsprachenlernern ausgerichteter einsprachiger deutscher Wörterbücher ist in den letzten 10 – 15 Jahren eine positive Entwicklung in Gang gekommen, wie auch Wiegand einschätzt (1998, X), unbenommen davon, wie die Kritik an den Wörterbüchern im Einzelnen auch ausfallen mag.

² Vgl. Roos (1975), Irsula Peña (1994), Bahns (1996), Gladysz (2003), Caro Cedillo (2004), Butina-Koller (2005), Reder (2006). Kookkurrenzanalysen auf der Basis der IDS-Korpora beziehen Steinbügl (2005) (wenn auch mit kritischem Impetus) und in systematischer Weise Hollós (2004) ein. Verwiesen sei hier noch auf die Arbeiten zu Kollokationen von Ulrich Heid und Kollegen, die an der Schnittstelle zwischen automatischer Sprachverarbeitung und Lexikografie anzusiedeln sind (vgl. z. B. Heid 2005).

deutschen Gegenwartssprache“ (WDG) kann – so Wiegands Einschätzung – „als die Pionierleistung der germanistischen Lexikographie im Bereich der großen mehrbändigen, einsprachigen Wörterbücher der deutschen Standardsprache der Gegenwart“ gelten (1990, 2130). Diese Bewertung betrifft nicht zuletzt die Einbeziehung von Kollokationen. Wer sich einen Eindruck vom Kollokationsreichtum des WDG verschaffen will, sollte beispielsweise das Lemma *Grund* nachschlagen. Agricolas „Wörter und Wendungen“ (Agricola) (1. Auflage 1962!) kann man mit Fug und Recht als ein frühes Kollokationswörterbuch oder besser gesagt, als ein Wörterbuch verfestigter Wortverbindungen des Deutschen bezeichnen. Die Vielfalt der Angaben ist bemerkenswert, wenn man bedenkt, dass damals noch keine deutschsprachigen elektronischen Korpora zur Verfügung standen.³

Stark geprägt ist die deutschsprachige Kollokationsforschung von der *Phraseologie* und ihrem Kernkonzept der ‚Idiomatizität‘, obwohl Kollokationen lange Zeit als phraseologisches Randphänomen angesehen wurden. Mittlerweile hat die Kollokation „einen festen Platz in der Phraseologie als Subkategorie (mit schwacher oder gänzlich fehlender semantischer Umdeutung), die den ‚Idiomen‘ als den semantisch markierten Phrasemen gegenübersteht.“ (Burger/Dobrovol’skij/Kühn/Norrick 2007b, 7). Mit dem Erscheinen der beiden HSK-Bände zur Phraseologie (Burger/Dobrovol’skij/Kühn/Norrick 2007a) und der 3. Auflage der „Einführung“ in die Phraseologie (Burger 2007) ist das Verhältnis zu Kollokationen als „nicht bzw. nur schwach-idiomatische Phraseologismen“ (ebd., 38) seitens der Phraseologen endgültig konzeptuell geklärt.

Die Arbeiten von Franz Josef Hausmann haben entscheidend zur Etablierung des Kollokationsbegriffs im deutschsprachigen Raum beigetragen. In richtungsweisenden Aufsätzen entwickelte Hausmann in den 80er Jahren die Eckpunkte seines zwar vielfach modifizierten, aber in den Grundzügen bis heute aktuellen Kollokationsmodells (vgl. Hausmann 1984, 1985, 1989, 2004). Siepmann bezeichnet ihn als Nestor der französischen Kollokationsforschung (2002, 240). Aber auch für die meisten deutschsprachigen bzw. mit der Kontrastsprache Deutsch arbeitenden lexikografischen, übersetzungswissenschaftlichen und fremdsprachendidaktischen Ansätze kann Hausmanns Modell als ein Leitkonzept angesehen werden. Holderbaum/Kornelius begründen die Geltungshoheit dieses Ansatzes so:

„Wenn die Typologie des gebundenen Sprachgebrauchs von Hausmann in die Übersetzungsforschung entlehnt wurde, so erklärt sich dies daraus, daß die Affinitätsbeziehung zwischen Basis und Kollokator die Bedingungen beim Übersetzen abbildet. [...] Es gibt komplexere, im Sinne der linguistischen Beschreibung leistungstärkere Modelle, [...] doch stehen diese der Übersetzungspraxis und den Bedingungen des Spracherwerbs eher fern.“ (2001, 537)

Die Kollokation ist eine „typische, spezifische und charakteristische Zweierkombination von Wörtern“ (Hausmann 1985, 118). Sie weist eine interne semantische Determiniertheit von Basis (semantisch autonom: *Ruhe, Glück*) und Kollokator (semantisch abhängig: *einkehren, unverschämt*) auf. Es handelt sich um konzeptuelle Relationen (*Ruhe – einkehren, unverschämt – Glück*), die aber auch in erweiterten Strukturen (z.B. *dringlichen Appell richten*) realisiert werden können. Teilidiome, z.B. *Bauklötze staunen*, werden subsumiert, Idiome dagegen explizit ausgeschlossen (alle Kollokationsbeispiele aus Hausmann 2004).

³ Vergleicht man den Agricola-Artikel ‚Grund‘ (1962, 252–253) mit aktuellen Korpusbefunden, so ist nach immerhin über 40 Jahren doch eine bemerkenswerte Stabilität in den syntagmatischen Umgebungen festzustellen. Hier wären sicherlich systematische Vergleiche sehr aufschlussreich.

Kollokationen sind sprachliche Phänomene, die primär für die Textproduktion, und zwar vorzugsweise aus der Sicht des Fremdsprachenlernalers, relevant sind. Daraus resultiert das Postulat der besonderen Stellung des Nomens als Basis, von der aus Lerner nach Kollokationen suchen. Nomina sollten demzufolge den Kernbestand eines Kollokationswörterbuchs ausmachen. Das *Oxford Collocations Dictionary* (OCD) sieht Hausmann als den vorbildlichen Prototyp eines Kollokationswörterbuchs an (2004, 311–312). In seinen frühen Arbeiten (vgl. 1985, 123–128) entwickelte er seine Ideen auch im Kontrast zur statistikbasierten ‚collocation analysis‘ des britischen Kontextualismus (vgl. Firth 1957, Sinclair 1991).⁴ Auf der 39. IDS-Jahrestagung 2003 reflektierte Hausmann noch einmal die verschiedenen terminologischen Paradigmen, um gleichzeitig Frieden anzubieten:

„Zu seinem Erstaunen stellt Momo fest, dass derzeit eine Art Krieg stattfindet, ein Terminologiekrieg, der Krieg um die Besetzung des linguistischen Terminus *Kollokation* [...] Der basisbezogene Kollokationsbegriff ist der engere, der merkmalsreichere, der elaboriertere, der genauere, der funktionalisiertere, der anwendungsbezogener, folglich der unverzichtbarere. [...] Ich nehme für mich in Anspruch, den Terminus von Firth fruchtbar weiterverarbeitet zu haben und betrachte mit dem Erscheinen des *Oxford Collocations Dictionary* den Krieg als gewonnen. Den Computerlinguisten dürfte es leicht fallen, ihren weiten Kollokationsbegriff mit einem anderen Terminus zu besetzen.“ (Hausmann 2004, 320–321)

Der Terminus ‚Kookkurrenz‘ hat gute Chancen, sich als ein weites Konzept für das statistisch signifikante Kovorkommen sprachlicher Einheiten im deutschsprachigen Raum durchzusetzen. Kookkurrenz ist das berechnete Phänomen. Kollokation verkörpert eine mögliche linguistische Sicht auf die Daten. (Steyer 2004b, 99; zu diesem Verhältnis vgl. u.a. auch Lehr 1996, Lemnitzer 1997, Feilke 2004, Duffner 2006, Lemnitzer/Zinsmeister 2006, Sailer 2007).⁵ Umfangreiche korpuslinguistische Forschungen in Bezug auf Kookkurrenzen werden am Institut für Deutsche Sprache in Mannheim geleistet. Dies betrifft sowohl die Entwicklung komplexer mathematischer Verfahren zur Erschließung sprachlicher Massendaten (vgl. u.a. Perkuhn 2007, Keibel/Belica 2007, KLb) als auch die linguistische Interpretation berechneter Korpusdaten im Spannungsfeld zwischen linguistischer Theoriebildung und praxisnahen Anwendungen (vgl. u.a. Steyer 2003, Steyer 2004b, Steyer/Lauer 2007, Brunner/Steyer 2007, Belica/Steyer i. Dr.; UWV). Ein zentrales Konzept für die linguistische Interpretation von Kovorkommen lexikalischer Einheiten auf der Basis von Frequenz- und Signifikanzberechnungen ist das Konzept der ‚usuellen Wortverbindungen‘ (UWV) (erstmal bei Steyer 2000, vgl. auch Steyer i. Dr.): Unter UWV verstehe ich kon-

⁴ Lehr (1996) und Lemnitzer (1997) veröffentlichten in den 90er Jahren einschlägige deutsche Studien zu Kollokationen unter systematischer Einbeziehung der Forschungen des britischen Kontextualismus und entsprechender korpusanalytischer Zugänge. Die darin entwickelten Ansätze zur Selektion von kookkurrierenden Einheiten aus Korpora sind nicht zuletzt auf das Ziel ausgerichtet, zu formal operationalisierbaren und maschinell weiter verarbeitbaren Resultaten für die Computerlexikografie zu kommen.

⁵ Die Benennungen wichtiger Tools zur Extraktion von Wortverbindungen und von entsprechenden Datenbanken wie die am IDS entwickelten stützen diese terminologische Entwicklung. Auch im Wortschatzportal der Universität Leipzig findet man heute u.a. automatisch berechnete ‚signifikante Kookkurrenzen‘ und ‚Mehrwortkookkurrenzen‘. Anzumerken wäre jedoch, dass ‚collocation‘ und ‚collocation analysis‘ in der englischsprachigen korpuslinguistischen Szene als Bezeichnung für ein signifikantes Miteinandervorkommen sprachlicher Zeichen bzw. für den entsprechenden korpuslinguistischen Zugang nach wie vor gebräuchlich ist.

ventionalisierte – zumindest binäre und minimal lexikalisch ausgefüllte – Muster des Sprachgebrauchs. Ihre Komponenten weisen eine auffällige statistische Affinität zueinander auf und sind in rekurrente, syntagmatische Strukturen eingebettet (usualisierte Syntagmen). UWV sind somit eher textuelle denn binär-lexikalische Zeichen (vgl. auch Bubenhofer 2008). Usuelle Wortverbindungen sind auf verschiedenen Abstraktionsebenen als holistische Einheiten interpretierbar. Die Fixiertheit kann sowohl zwischen einzelnen Wortformen sprachlicher Ausdrücke als auch als Relation zwischen lexikalischen Einheiten und Ausdrucksklassen manifest werden:⁶

- (1)
auf Grund
Grund genug
in Grund und Boden
aus welchen Gründen auch immer
aus [Adjektiv_{negativ-bewertend}] (z.B. *fadenscheinigen, unerfindlichen*) *Gründen*
im Grunde [Possessivpronomen/Artikel] (z.B. *meines/deines/des*) [Nomen_{Körper/Mensch}] (z.B. *Herzens/ Seele/ Wesens*)

Diese Fixiertheit ist aber nicht nur eng semantisch zu interpretieren, z.B. *kompositioneller* vs. *nicht-kompositioneller* oder *idiomatischer* vs. *nicht-idiomatischer Status*, sondern primär pragmatisch. Es handelt sich um eine Fixiertheit „innerhalb konventionaler Gebrauchskonstellationen.“ (Feilke 2004, 47). Die Kollokationsperspektive fokussiert einen Teilaspekt usueller Wortverbindungen, nämlich die interne Bestimmtheitsrelation zwischen autosemantischen Hauptkomponenten usueller Wortverbindungen (Basis und Kollokator). Im Mittelpunkt steht hier die konzeptuelle Affinität der Elemente zueinander. Kollokationen haben als Teilaspekt von UWV folgerichtig auch das Kriterium der Rekurrenz, ermittelt durch statistische Signifikanzmessungen, zu erfüllen (vgl. Siepmann 2002, 254; 2005, 411). Auf Übergangsbereiche zwischen Kollokationen und usuellen Wortverbindungen werde ich in Kapitel 5 eingehen. Im Folgenden wird der Terminus ‚Kollokation‘ im Sinne von Hausmann verwendet.

3 Neues Kollokationswissen. Korpora, Ressourcen, Portale

Schon heute kann (oder könnte) jeder Lexikograf auf mächtige online abrufbare lexikalische korpusbasierte Ressourcen für das Deutsche zurückgreifen.⁷ Das Deutsche Referenzkorpus *DeReKo*, beheimatet am Institut für Deutsche Sprache, bildet „mit über 3.2 Milli-

⁶ Das UWV-Konzept hat Schnittstellen zu Ansätzen der *Konstruktionsgrammatik* (vgl. Fischer/Stefanowitsch 2006a). Allerdings werden lexikalisch nicht ausgefüllte Konstruktionen explizit ausgeschlossen (Fischer/Stefanowitsch 2006b; vgl. Deppermann 2006, 48–49).

⁷ Hier werden nur die wichtigsten öffentlich und kostenlos für jeden Nutzer weltweit zugänglichen Ressourcen, bezogen auf das Deutsche, aufgeführt (ein Vergleich in Duffner/Näf 2006). Einen umfassenden Überblick auch über andere, hier nicht genannte korpuslinguistische Ressourcen und Methoden gibt Sailer 2007. Zu erwähnen ist die wachsende Zahl von Arbeiten, die das Word Wide Web als Ressource und Quelle nutzen bzw. mehrere Zugänge (Korpora, Web, elektronische

arden Wörtern die weltweit größte linguistisch motivierte Sammlung elektronischer Korpora mit geschriebenen deutschsprachigen Texten aus der Gegenwart und der neueren Vergangenheit“ (KLa). *DeReKo* kann somit als größter Ausschnitt der deutschen Gegenwartssprache gelten und ist auf Grund der quantitativen Dimension eine ernstzunehmende Basis für statistische Signifikanzberechnungen des Deutschen. Das Korpus ist mit Hilfe des Korpusrecherche- und Analysesystems COSMAS durchsuch- und analysierbar (CII). Das zweite große online nutzbare elektronische Korpus des Deutschen ist das Korpus des Digitalen Wörterbuchs der deutschen Sprache des 20. Jahrhunderts an der Berlin-Brandenburgischen Akademie (DWDS). Öffentlich zugänglich sind ein „zeitlich und nach Textsorten ausgewogenes Corpus des gesamten 20. Jahrhunderts. Umfang: 100 Millionen Textwörter (tokens)“ (ebd.) und diverse Spezialkorpora. Auch für das DWDS-Korpus steht eine linguistische Suchmaschine zur Verfügung (vgl. Fellbaum 2007).

Bei der Erschließung und Auswertung dieser Korpora ist zu unterscheiden, ob es sich um Zugänge zu rein automatisch aufbereiteten Sprachdaten einschließlich der entsprechenden Analyseverfahren auf der einen Seite handelt oder um die Onlinenutzung neu erarbeiteter lexikografischer Ergebnisse unter Ausnutzung dieser Ressourcen auf der anderen, z. B. in lexikografischen Informationssystemen (vgl. dazu Engelberg/Lemnitzer 2001; Müller-Spitzer 2007).

Zugang zu automatischen Kookkurrenzdaten. Das *Wortschatzportal der Universität Leipzig* (Wortschatz) enthält neben vielen anderen Informationen auch automatisch ermittelte Partnerwörter mit entsprechenden Korpusstellen von mehr als 9 Millionen Wörtern und Wortgruppen (direkte Nachbarn, Kookkurrenzen usw.). Man kann jedoch nicht selbst interaktiv und dynamisch im Korpus recherchieren und entsprechende Kookkurrenzanalysen vornehmen. Kritisch anzumerken wäre hier die nach wie vor mangelhafte Dokumentation des zu Grunde liegenden Korpus und der angewendeten Methoden im Online-Auftritt. Methodische Transparenz ist aber für linguistische und lexikografische Bewertungen der Resultate unabdingbar. Auch im DWDS lassen sich Kookkurrenzpartner von Suchobjekten (dort „automatisch berechnete Kollokationen“) nach verschiedenen statistischen Maßen sortiert abrufen und mit den entsprechenden Kontextzeilen anzeigen, allerdings auf einer sehr viel kleineren Korpusbasis als beispielsweise das IDS-Korpus. Dort ist es auch nicht möglich, individuelle Parameter zu definieren, so dass man u. a. obligatorisch die Grundform eines Kookkurrenzpartners erhält. Das IDS bietet derzeit zwei Zugänge zu automatisch berechneten Kookkurrenzen. Die Kookkurrenzdatenbank CCDB (Belica 2001–2007) beinhaltet – öffentlich zugänglich – automatisch berechnete Kookkurrenzprofile zu ca. 220 000 Wörtern, in verschiedenen Profilen und unter verschiedenen Aspekten visualisiert (vgl. Keibel/Belica 2007). Die u. a. in COSMAS II implementierte *statistische Kookkurrenzanalyse* (Belica 1995) gehört zu den komplexesten automatischen Verfahren zur Selektion von signifikanten Wortkombinationen, die für deutsche Korpora online und kostenlos zur Verfügung stehen. Sie berechnet und visualisiert nicht nur binäre Kookkurrenzen wie viele andere Tools, sondern ermittelt vor allem auch die rekurrenten syntagmatischen Strukturen, in die signifikante Wortkombinationen eingebettet sind, und fasst die jeweiligen Originaltextauschnitte (KWICs und Korpusstellen) in Signifikanzcluster zusammen. Ein großer Vorteil

Wörterbücher) miteinander kombinieren: Stellvertretend hierfür sei das von Peter Dürco verantwortete Projekt „Kontrastive Kollokationsforschung Slowakisch-Deutsch“ genannt.

besteht des Weiteren darin, dass man diese Analysen nach individuellen Parametern (z.B. Kookkurrenzspanne oder Kontextspezifik) und auf einer variablen Korpusbasis (entweder auf dem aktuellsten Stand aller öffentlich zugänglichen IDS-Korpora oder auf der Basis eines individuell komponierten – virtuellen – Teilkorpus) dynamisch vornehmen kann (vgl. KLb). Allerdings erfordert die sinnvolle Anwendung der Kookkurrenzanalyse eine gewisse Methodenkompetenz bzw. Einarbeitungszeit.⁸

Das größte Vorhaben im Bereich *elektronischer wissenschaftlicher Lexikografie* unter Ausnutzung umfangreicher Korpora ist das Online-Wortschatz-Informationssystem Deutsch (OWID) am IDS, das mehrere Komponenten beinhaltet (derzeit Bedeutungswörterbuch *ele-xiko*, Neologismenwörterbuch, feste Wortverbindungen, Diskurswörterbuch) (vgl. Haß 2005; Klosa 2008). Kookkurrenzen werden in OWID zum einen als heuristischer Zugang zu usu-ellen Bedeutungen der Lemmata genutzt. Zum anderen werden sie auf verschiedene Weise als lexikografisches Ergebnis aufbereitet: als semantische Mitspielerangaben (zum Teil handelt sich um Kollokationen, zum Teil um eher thematische oder framebezogene Partner) und als typische Verwendungsmuster (phrasale Angaben auf der Basis von syntagmatischen Mustern) in *ele-xiko* sowie in gesonderten Mehrwortartikeln als eigener Beschreibungsgegenstand (vgl. Steyer/Lauer/Brunner 2008). Seit Kurzem sind alle *ele-xiko*-Stichwörter mit der IDS-Kookkurrenzdatenbank verlinkt. Auch DWDS (WDG → Korpusrecherche → Kollokationsanalyse) und Wortschatz (automatische Wortinformationen → Dornseiff) vernetzen auf diese Weise automatische Ressourcen und lexikografische Ergebnisse.

4 Korpuslinguistische Fundierung? Kollokationen in deutschen Wörterbüchern

Wie bereits erwähnt, gibt es noch nicht *das* Kollokationswörterbuch des Deutschen. Hausmann sieht im Stilduden (Duden 2) zwar erste Ansätze für ein solches Spezialwörterbuch, moniert aber, dass Duden 2 „auf halbem Wege stehen bleibt“ (2004, 312). Schemanns Idiomatik (Schemann 1993) und der Redewendungsuden (Duden 11) sind zwei wichtige Referenzwerke für die Dokumentation von Wortverbindungen des Deutschen. Sie enthalten jedoch neben Kollokationen auch alle anderen Wortbindungstypen (vor allem Idiome) und können deswegen nicht als reine ‚Kollokationswörterbücher‘ angesehen werden. Ganz so defizitär scheint die Situation dennoch nicht, denn es finden sich über die wenigen eben genannten Spezialwörterbücher hinaus auch in allen einsprachigen Wörterbüchern des Deutschen (sowohl in Bedeutungs- als auch in Lernerwörterbüchern) Kollokationen in beachtlicher Zahl und an sehr verschiedenen Stellen der Mikrostruktur wieder. Im Rahmen groß angelegter Wörterbuchanalysen (vgl. Wiegand 1998, 2002, 2003, 2005) wurde auch die Behandlung von Kollokationen detailliert untersucht: LGWDAF von Lehr 1998; DWDAF von Köster/Neubauer 2002 und DGWDS von Schafroth 2003 (vgl. auch Lehr 1997, Reder 2006, 167–188). Für unsere Zwecke sind darüber hinaus die Lernerwörterbücher PGW-DaF und HWDaF sowie die einsprachigen Bedeutungswörterbücher DDUW und WDW

⁸ Die Aktivitäten meines Projektes ‚Usuelle Wortverbindungen‘ zur Vermittlung von linguistischer Methodenkompetenz, zu entsprechenden Transferleistungen und Arbeitskontakten sind auf der Projekthomepage dokumentiert (UWV).

von Interesse, da sie 2006 und 2007 zum Teil neu bzw. in neuen Auflagen erschienen sind und deshalb eine korpusmethodische Fundierung erwarten lassen. Aus korpuslinguistischer Perspektive ist zunächst folgende Frage zu beantworten: Sind in den Wörterbüchern neue methodische Ansätze in Bezug auf das *Prinzip der Korpusbasiertheit* im Allgemeinen und auf eine korpusbasierte Selektion und Systematisierung von Kollokationen und Beispielen im Besonderen zu erkennen? Dass die Wörterbuchredaktionen von heute mit elektronischen Korpora z.B. zum Abgleich ihrer Stichwortlisten arbeiten, weiß man aus Fachveröffentlichungen. Die entsprechenden Informationen dazu sollten aber auch in den Außentexten der Wörterbücher dokumentiert sein: Im Vorwort des DGWDS werden bereits 1999 elektronische Korpora als Quellenmaterial erwähnt ebenso wie Recherchemöglichkeiten im Internet. Das DDUW geht auch auf die Vielfalt benutzter Quellen ein, u.a. auf das seit 2001 kontinuierlich ausgebauten Duden-Korpus, das derzeit „700 Millionen mit sprachbezogenen Informationen versehene Wortformen“ enthält (13). Im LGWDaF findet sich der Vermerk, dass ein „umfangreiches Korpus eine ergiebige Quelle für authentische Beispiele“ sei. Aus didaktischen Gründen hätte man aber „darauf verzichtet, nur solche Beispiele anzuführen, die einem Korpus entstammen.“ (6). Um welche Korpora es sich handelt, ob hauseigene oder externe wie die IDS-Korpora, erfährt man hier nicht. Im Vorwort zur Neuausgabe des WDW 2006 ist zu lesen, dass auf der Basis des *WAHRIG Textkorpus digital* „10.000 neue Stichwörter recherchiert und der Häufigkeit ihres Vorkommens entsprechend ausgewählt“ (WDW, 1) wurden. In den aktuellen Wörterbuchauflagen wird also durchaus – wenn auch oft noch eher versteckt – auf die Nutzung von Korpora verwiesen. Nirgendwo wird jedoch dargelegt, auf welchen methodischen Wegen die Korpusbefunde dann als Evidenzen für Gebrauchstypik und Usualität genutzt werden. Diese nicht vorhandene Transparenz in Fragen der Empirie ist umso unverständlicher, als nach meiner Kenntnis beispielsweise auch in der Duden-Redaktion inzwischen automatische Methoden zur Kollokationsextraktion entwickelt (und angewendet?) werden. Andere Kookkurrenzanalysen und lexikalische Ressourcen stehen – wie oben erläutert – ja ohnehin teils seit mindestens 10 Jahren online zur Verfügung. Eine augenfällige Dokumentation der Korpusbasis und eine Reflexion der Methodik der Wissensextraktion sind aber m.E. für ein modernes Wörterbuch des 21. Jahrhunderts unabdingbar.⁹ Jedem lexikografischen Arbeitsprozess sollten heute Korpusanalysen und im Idealfall auch Kookkurrenzanalysen vorgelagert sein. Das Kollokationswörterbuch der Zukunft kann eigentlich nur noch auf Basis statistischer Berechnungen erarbeitet werden. Sicherlich erfassen auch statistische Korpusanalysen nicht alle potenziellen Kollokate einer Basis. Aber diejenigen, die ermittelt wurden, sind bei entsprechender quantitativer Korpusbasis von *gesicherter Usualität*. Und diejenigen Kollokationen, die kompetenzbasiert abgeleitet wurden und wirklich gebräuchlich sind, finden sich in der Regel auch in den Kookkurrenzergebnissen wieder: So entdeckt auch die IDS-Kookkurrenzanalyse, dass eine Ente typischerweise *watschelt*, *schnattert* oder *schwimmt* (Ente-Beispiel aus Hollós

⁹ Wie man auch korpusmethodische Informationen in einem Wörterbuch verarbeiten kann, zeigt das *New Oxford Dictionary of English* (NODE). Dort wird nicht nur ausführlich auf die Korpusbasis eingegangen, sondern auch die Gewinnung von Evidenzen und Beispielen über Konkordanzen und KWICs illustriert (NODE, XII–XIII.); vgl. auch die Außentexte im *Oxford Collocations Dictionary* (OCD 2002, viii).

2004) oder dass ein Anker *gelichtet* bzw. *(aus)geworfen* wird (Anker-Beispiel aus Hausmann 2004.). Der *himmelweite Unterschied* (ebd.) ist in der Tat statistisch sehr auffällig. Die Kookkurrenzanalyse zeigt aber darüber hinaus, dass Unterschiede – statistisch signifikant – auch *gravierend, gewaltig, fundamental, prinzipiell, signifikant, deutlich, erheblich, grundsätzlich, riesig, gewichtig, markant, eklatant, augenfällig, krass, essentiell, elementar, riesengroß, spürbar, gehörig, ungeheuer, enorm, nennenswert, tief greifend, frapierend, beträchtlich, drastisch, immens* sein können. Und man erfährt mehr: Der *himmelweite Unterschied* kommt sehr viel häufiger vor als *himmelweite Unterschiede*. Es gibt außer *Unterschied* kein weiteres Nomen, dass signifikant auffällig mit *himmelweit* kollokiert. In der Umgebung des Adjektivs *himmelweit* ist nur noch ein weiteres Syntagma von statistischer Auffälligkeit, nämlich *himmelweit entfernt sein*. Im Gegensatz dazu ist *Unterschied* beispielsweise im Kookkurrenzprofil von *eklatant* nur eines von vielen affinen Nomen wie *Mangel, Verstoß, Fehler, Schwächen*. Und so kann man die Kollokationsnetze auf der Basis von Kookkurrenzen weiter verknüpfen.

Die bisherigen Kritiken der Wörterbuchforschung an der Kollokationspraxis in deutschen Wörterbüchern im Einzelnen beziehen sich in der Regel auf folgende Aspekte:

- a. Auswahl und Auffindbarkeit von Kollokationen in den Wörterbüchern (u.a. unter welchem Stichwort: Basis oder Kollokator?)
- b. Abgrenzung zu freien Wortverbindungen einerseits und anderen Wortverbindungstypen wie Idiomen andererseits
- c. Stellung in der Mikrostruktur
- d. Geltungsbereich von Kollokationen (Restriktionen)

Die Punkte a. und b. sind in der angegebenen Literatur ausführlich besprochen worden (z. B. bei Lehr 1998, 269–274; Köster/Neubauer 2002, 288–289, 299–303; Schaefroth 2003, 399–403) und werden hier nicht weiter behandelt. Bezüglich der Auswahl der Kollokationen sei nur angemerkt, dass sich in den Einträgen der untersuchten Wörterbücher selbst keine klaren Indizien dafür finden, ob auf der Basis von Korpora und möglicherweise sogar mit Kookkurrenzanalysen gearbeitet wurde. So sind die Kollokationen zum Stichwort *Grund* zum großen Teil schon bei Agricola und im WDG zu finden, natürlich mit Abwandlungen und Ergänzungen. Das heißt nun nicht, dass diese Kollokationen heute in jedem Falle untypisch oder veraltet wären. Es lässt sich in den aktuellen Wörterbucheinträgen nur nicht erkennen, ob und inwieweit mittlerweile Korpusbefunde eingeflossen sind. Auf den zentralen Aspekt des Geltungsbereichs von Kollokationen (d.) wird in einem gesonderten Kapitel (5.) vor dem Hintergrund kookkurrenzbasierter Vorgehens ausführlich eingegangen.

Ein immer wieder diskutierter Kritikpunkt bezieht sich auf die *Stellung von Kollokationen in der Mikrostruktur* (c.): Wortverbindungen und Kollokationen werden zwar in allen Außentexten explizit erwähnt. Zumeist wird auch die Systematik ihrer Darstellung erklärt. Das ändert aber nichts am eher unauffälligen Erscheinungsbild von Kollokationen in den Wörterbuchartikeln selbst. Nach wie vor haben sie nicht den zentralen Stellenwert in der Mikrostruktur, der ihnen gebührt. Durch die Vermischung mit anderen lexikografischen Inhalten sind sie immer noch zu unscheinbar, um ein wirkliches „Kollokationsbewußtsein“ beim Wörterbuchnutzer zu erreichen (Köster/Neubauer 2002, 304; vgl. schon Viehweger 1989). Kollokationsbewusstsein meint hier die Sensibilisierung dafür, dass Kollokationen zentrale Konstituenten des Sprachgebrauchs darstellen, denen mindestens genauso viel Aufmerksamkeit gebührt wie der Rechtschreibung oder den grammatischen Regeln. Kolloka-

tionen werden so immer noch zu wenig als eigene Qualität der Sprachbeherrschung wahrgenommen werden, an der sich angemessener Sprachgebrauch in erster Linie misst.¹⁰

Kollokationen im engen bzw. Kookkurrenzen im weiten Sinne bilden darüber hinaus zentrale Zugänge zu usuellen Bedeutungen eines Lemmas (vgl. z. B. Firth 1957, Sinclair 1991, Heringer 1999, Stubbs 2001, Blumenthal 2006). Warum sollte sich diese Heuristik in Bedeutungswörterbüchern nicht auch für den Nutzer explizit nachvollziehen lassen (vgl. Konerding 1998)? Hausmann unterscheidet in diesem Sinne zwischen Wörtern, die semantisch autonom erklärbar sind und solchen, für deren Definition Kontextpartner unverzichtbar sind (1997, 173). Letztere müssten auch in den Bedeutungswörterbüchern explizit über möglichst viele und möglichst systematische Kollokationen (bzw. allgemeiner: Kookkurrenzen) erklärt werden (vgl. Beispiel *abverlangen* Steyer 2004b, 105–106). Aber auch eher thematisch bzw. framebezogene Kookkurrenzpartner könnten einen Platz in der Bedeutungs-erklärung finden. (vgl. dazu das Beispiel *Wasser* in Steyer 2004b, 104–105).

5 Erklärungsansätze und Visionen. Kookkurrenzen als Basis für Kollokationswörterbücher

Im folgenden Punkt geht es vor allem um das Erklärungspotenzial der Kookkurrenzperspektive für ein Kollokationswörterbuch, sowohl unter einer weiten Forschungsperspektive als auch unter dem Blickwinkel des derzeit ‚konkret Machbaren‘. ‚Kollokationswörterbuch‘ soll hier aber nicht nur als ein Produktionswörterbuch verstanden werden. Wir haben ebenso lexikografische Produkte im Blick, die für gehobenes bis muttersprachliches Kompetenzniveau (z. B. für Übersetzer, DaF-Lehrer, für professionelle Textproduzenten) auch in rezeptiver Hinsicht sinnvoll sein können. Bei solchen Wörterbüchern geht es vor allem um Sprachreichtum und um die Erschließung der kontextuellen Feinheiten des distinktiven Gebrauchs sprachlicher Ausdrücke über Korpusdaten. Das Erklärungspotenzial der Interpretation von Kookkurrenzdaten ist in jedem Falle beträchtlich, ungeachtet dessen, mit welcher Elaboriertheit, Systematik oder didaktischen Strenge diese Ergebnisse dann in die einzelnen Wörterbücher einfließen. Es bezieht sich insbesondere auf den *Kollokationsreichtum* und die *Kollokationsselektion*, auf *Kollokationsrestriktionen* und *kontextuelle Gegebenheiten des Gebrauchs* sowie auf die Rekonstruktion von *abstrakten Bildungsmustern*. Bezüglich des *Kollokationsreichtums* genügt ein Blick in Kookkurrenzlisten. Beispielsweise ließe sich eine solche Liste von expressiven Adjektiven wie sie im Zusammenhang mit *Unterschied* gegeben wurde, nur schwer kompetenzbasiert mit Hilfe von Belegsammlungen oder mittels Befragungen bzw. Assoziationstests erstellen. Die Frage, welche dieser Kollokationen in ein Wörterbuch aufgenommen werden, kann nur mit Hilfe folgender Kriterien beantwortet werden: statistische Signifikanz, Nutzungssituation (Rezeption/Produktion;

¹⁰ In Bezug auf Lernerwörterbücher hebt Herbst u.a. die Kollokationspraxis des LGWDAF hervor und schreibt ihr in Bezug auf die systematische Erfassung von Kollokationen gar Vorbildfunktion für die britische Lernerlexikografie zu (1998: 30). Lehr diskutiert dagegen einige Kritikpunkte (1998). Auf einen zentralen Punkt wird in 5. eingegangen.

Kompetenzniveau der Adressaten; Relevanz im Vergleich zur Kontrastsprache) und Art des Präsentationsmediums (gedruckt oder elektronisch).

Eine zentrale Kritik an der bisherigen Kollokationspraxis in deutschen Wörterbüchern bezieht sich darauf, dass aus den Kollokationsangaben zumeist nicht auf den *Typikalitätsgrad* der angegebenen Kollokatoren zu schließen ist. Es fehle eine präzise Darlegung der jeweiligen Kollokationsgrenzen (Lehr 1998, 278). In der Regel kann der Nutzer keine oder nur wenig Rückschlüsse auf den Geltungsbereich von Kollokationen ziehen. Allerdings ist dieses Manko keinesfalls lexikografischer Oberflächlichkeit geschuldet. Auch die Kritiker haben keine wirklichen alternativen Vorschläge entwickelt. Und das hat objektive Gründe: Ohne korpusanalytisches Instrumentarium kann man Regularitäten in der präferierten Kombinierbarkeit lexikalischer Einheiten (oder anders herum: Ausschlussbedingungen für Kollokatoren) nur mit Hilfe des Kriteriums der *besonderen Auffälligkeit*, sprich einer irgendwie gearteten Beeinträchtigung der Transparenz der Wortkombination, herleiten. Oder der Geltungsbereich wird durch das Ausschließen ‚unmöglicher Kollokate‘ abgesteckt. Dies ist aber – darin ist sich die Forschung im Grunde einig – nur im kontrastiven Vergleich sinnvoll: L1 vs. L2. Der vielfach diskutierte Ausweg besteht darin, nach Verallgemeinerungen auf der konzeptuellen Ebene zu suchen, auf die sich einzelne lexikalische Einheiten zurückführen lassen. Laut Scherfer beruhen Kollokationsrestriktionen darauf, „dass eine Sprache kognitive Synonyme (Cruse 1986: 88–92 u. 270–285) für das abhängige Mitglied der Wortverbindung zur Verfügung hat, aus denen im Hinblick auf das andere ausgewählt werden muss.“ (2001, 233). Dieser Auswahlprozess jedoch folgt keinem Regelsystem, sondern vollzieht sich qua Konvention (in diesem Sinne z.B. Siepmann 2002, Feilke 2004, Brunner/Steyer 2007). Feilke spricht von Selektions- und Kombinationspräferenzen aus einem Spektrum von Konstruktionsmöglichkeiten, die durch Konventionalisierung verbindlich geworden seien (1998, 74). So werden die Kollokationen *fadenscheinig* – *Grund* und *durchsichtig* – *Grund* im IDS-Korpus fast immer im Plural verwendet (z.B. *aus/unter fadenscheinigen Gründen*; *aus durchsichtigen Gründen*). *Durchsichtig* – *Grund* ist darüber hinaus beschränkt auf den Gebrauch mit der Präposition *aus*: **unter/mit durchsichtigen Gründen* wird so gut wie nie verwendet (Eine entsprechende Google-Abfrage bestätigt diesen Korpusbefund). Natürlich könnte man sich Verwendungskontexte mit Syntagmen wie *aus einem fadenscheinigen Grund*; *aus einem durchsichtigen Grund* bzw. *mit einem durchsichtigen Grund* vorstellen. Diese Verwendungen wären regulär und in einer bestimmten Kommunikationssituation auch angemessen. Nur sagt man das selten. Solche Präferenzsetzungen sind nicht a priori und regelgeleitet synthetisierbar bzw. deubar. Dies lässt sich nur a posteriori erschließen, auf der Basis empirischer Befunde. Gemeint sind dabei keine Einzelbefunde, sondern ‚kollektivierte‘ Massendaten. Exhaustive statistische Kookkurrenzberechnungen liefern die entscheidenden Evidenzen in Bezug auf die Wahrscheinlichkeit der Kombinatorik von Kollokationskomponenten und ihrer ko- und kontextuellen Gebundenheit. Gibt man solche Begrenzungen im Gebrauch im Wörterbuch an, bedeutet das also eigentlich nur: *Diese Verwendungsweisen sind besonders typisch* (‚Du bist mit ihnen auf der sicheren Seite‘) und nicht: Andere denkbare Formen sind per se ausgeschlossen.

Betrachtet man Kookkurrenzen in Hinblick auf signifikante lexikalische Kollokatoren, also das Verhältnis von Basis und Kollokator, lassen sich zunächst grob drei ‚Bindungsqualitäten‘ unterscheiden:

- a. Exklusiver Status eines Kollokators
- b. Dominanz eines Kollokators
- c. Gruppenzugehörigkeit eines Kollokators.

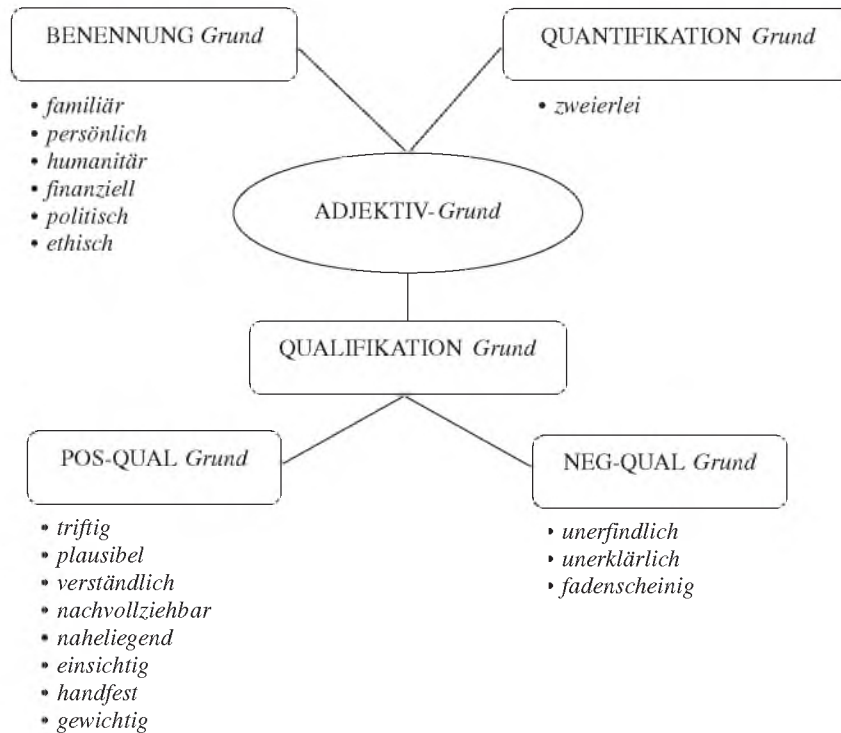
Es geht also darum zu überprüfen, ob es neben einem als Kollokator identifizierten Kookkurrenzpartner noch andere Vertreter ähnlicher Art gibt, die zum selben Kollokationsmuster in Bezug auf die Basis gehören. Kollokationsmuster sind aber nicht nur als wortartenbasierte Strukturtypen zu verstehen, etwa alle adjektivischen Kollokatoren zu einem Nomen, sondern sie untergliedern sich in semantisch und pragmatisch motivierte Ausdrucksklassen. Signifikante Kollokatoren bilden dabei die konkreten lexikalischen Füllungen, die zum gegebenen Zeitpunkt auf einer gegebenen Korpusbasis mit gegebenen Parametern berechnet wurden. Die lexikalischen Füller sind im schematischen Rahmen des Kollokationsmusters austauschbar (vgl. auch ‚Schematisierung‘ bei Eichinger 2004, 11). Die Varianz lexikalischer Füller kann schon dann zu Tage treten, wenn sich die Parameter der Analyse ändern, z.B. in Bezug auf die angesetzte Kollokationsspanne. Das Kollokationsmuster sollte dabei jedoch nicht angetastet werden. Das bedeutet, dass man bei der Interpretation Kollokatoren dahingehend qualifiziert, ob sie als typische Vertreter für eine ganze Ausdrucksklasse eines Paradigmas fungieren, mit der eine Basis Kollokationsmuster konstituiert.

Fall a. *Exklusiver Status eines Kollokators*: Es gibt keine weiteren lexikalischen Vertreter ähnlicher Art in den berechneten Kookkurrenzen dieses Bezugsworts, die dem Kollokationsmuster zuzurechnen sind: *im Grunde genommen*, *auf Grund laufen*, *Grund und Boden*, *in Grund und Boden*. Das Kriterium ist hier nicht der vermeintliche Grad der Idiomatizität, sondern die lexikalische Exklusivität in Bezug auf das Ausgangswort.

Fall b. *Dominanz eines Kollokators*: Ein Kookkurrenzpartner ist besonders signifikant. Nur wenige andere Kookkurrenzpartner lassen sich als Vertreter ähnlicher Art identifizieren. So ist z.B. *Grund genug* in genau dieser lexikalisch-syntaktischen Struktur hochgradig signifikant und fest. Es gibt nur noch zwei – sehr viel seltenere – Vorkommen, die demselben Schema [Nomen_{Begründung}+*genug*] folgen, nämlich *Anlass (genug)* und *Motivation (genug)*. Empirisch validieren kann man solche stark lexikalisierten Affinitäten durch Reziprokanalysen des Kookkurrenzverhaltens der Kollokatoren, hier die Kookkurrenzanalyse des Adverbs *genug*. Es zeigt sich, dass *Grund* das mit Abstand affinste Nomen im unmittelbaren Vorfeld von *genug* ist und nur noch *Anlass* und *Motivation* statistisch auffällige Nomen mit sehr viel geringerer Frequenz sind.

Fall c. *Gruppenzugehörigkeit eines Kollokators*: In den meisten Fällen lassen sich die berechneten Kollokatoren jedoch zu Ausdrucksklassen zusammenfassen, z.B. zu adjektivischen Kollokatorengruppen mit bestimmten Merkmalen zur Lemmabasis *Grund*:

(2)



Eine solche Systematik, hier von Adjektiven, die einen Grund benennen *oder* quantifizieren *oder* positiv – negativ – changierend bewerten, lässt sich nur a posteriori auf der Grundlage systematischer Datenanalysen herstellen. Offen ist, welche dieser Abstraktionsebenen einen sinnvollen Zugriff aus Nutzerperspektive ermöglichen. Dazu sind noch umfangreiche, auch didaktische Forschungen nötig.

Bisher wurde die binäre lexikalische Restriktion als konzeptuelle Relation zwischen autosemantischen Wortschatzeinheiten betrachtet. Diese Zweigliedrigkeit von Kollokationen ist in der Forschung inzwischen vielfach relativiert worden, nicht zuletzt durch Hausmann selbst, der auch Triplestrukturen für Kollokationen einschließt (2004, 316). Siepmann schlägt ganz in diesem Sinne eine Erweiterung des Kollokationsbegriffs auf ‚long-distance collocations‘ (2005, 431) vor. Analysiert man Kookkurrenzstrukturen, erscheint diese Binariät in der Tat nur noch als ein Konstrukt. Fast immer sind Kollokationen in rekurrente, syntagmatische Strukturen eingebettet. Das bedeutet, dass auch die einbettende Umgebung von Basis und Kollokator durch weitere signifikante Auffälligkeiten im Gebrauch restringiert ist. Dies gilt nicht nur für präpositionale Anschlüsse und andere eng grammatische Merkmale. Es gilt vor allem für komplexe syntagmatische Einbettungen (*rekurrente Kotextmuster*), die den eigentlichen usuellen Gebrauch einer Kollokation konstituieren. Diese sind nur über automatische Analyseverfahren verifizierbar. Es geht also um die „textuelle Verankerung“ einer Kollokation (Schafroth 2003, 405). Aus korpusanalytischer Sicht bedeutet dies nicht die Verankerung im *Einzeltext*, sondern die Einbettung in wiederkehrende, sequenzielle

Muster *über viele Texte hinweg*. Hier sind wir bei usuellen Wortverbindungen angelangt. So wird die Verb-Nomen-Kollokation *sehen – Grund* fast durchgängig in Negationsstrukturen vom Typ *keinen Grund sehen* verwendet. Eine typische sprachliche Realisierung der Adjektiv-Nomen-Kollokation *gut – Grund* ist die satzwertige Ellipse *Und das aus gutem Grund!* Die Nomen-Verb-Kollokation *haben – Grund* wird häufig in der festen syntaktischen Struktur *allen Grund haben* realisiert. Auch unser Kollokationsmuster [Adjektiv] – *Grund* weist typische Kotextualisierungen auf: Das signifikanteste Syntagma ist *aus* [Adjektiv] – *Grund*. Interpretiert man Kollokationen als (ko-)textuelles Zeichen, ist ein Kollokationsmuster mit der Basis als Nennform *Grund*_{Lemma} jedoch eigentlich unzulässig: Basen und Kollokatoren kommen nicht in ihrer abstrahierten Grundform im Text vor, sondern als konkrete Wortformen. Das bedeutet, dass das Kookkurrenzverhalten daraufhin zu überprüfen wäre, ob die affinen Verhältnisse überhaupt für das ganze Flexionsparadigma verallgemeinerbar sind oder ob nicht auch *wortformenspezifische Verbindbarkeiten* auftreten (vgl. Tognini-Bonelli 2001, 92–98).¹¹ Vergleichen wir also das Kookkurrenzverhalten von *Grund*_{Singular} und *Gründen*: Es gibt Kollokationen, die auch bei einer Varianz in den Flexionsmerkmalen wie unterschiedlicher Kasus oder Numerus der Basis *Grund*_{Lemma} immer hervortreten. Das Adjektiv *plausibel* verbindet sich mit mehreren Flexionsformen von *Grund*_{Lemma}: *aus einem plausiblen Grund; aus plausiblen Gründen; plausible Gründe*. Andere dagegen sind an bestimmte Wortformen gebunden: So gibt es im Kookkurrenzprofil von *Grund*_{Singular} keine statistisch signifikanten adjektivischen Partner der Gruppe [Benennung] wie *finanziell, politisch oder persönlich*. Diese Adjektive kommen signifikant häufig nur in Verbindung mit den Pluralformen *Gründen* und *Gründe* vor. Das häufigste syntagmatische Muster ist *aus* [Adjektiv_{Benennung}] *Gründen* (z.B. in der Verwendung *aus persönlichen Gründen von seinem Amt zurücktreten, absagen; aus finanziellen Gründen etw. doch nicht investieren* usw.). Diese Restriktion ist eine Restriktion im Gebrauch und funktional zu erklären: Mit der Pluralform wird eine gewisse Vagheit ausgedrückt, die es dem Sprecher ermöglicht, die Gründe nicht näher spezifizieren zu müssen, obwohl sie auf Grund der Adjektivwahl als expliziert erscheinen. Interessant ist nun, dass auch in fast allen Wörterbucheinträgen diese Pluralrestriktion ‚durchschimmert‘, indem die Lexikografen intuitiv dem Usus gemäße sprachliche Formen von Kollokationsbeispielen wie *aus familiären Gründen absagen* vs. *aus gutem Grund* angegeben haben. Nur lässt sich eben nicht erkennen, ob z.B. *familiär* ein dominantes Exemplar eines Kollokatoren ist oder ob man Verallgemeinerung auf eine spezifische Adjektivgruppe treffen könnte, die dieser Numerusrestriktion unterliegen.

Es sei noch einmal zusammengefasst: Kollokationsmuster fokussieren die konzeptuelle Austauschbarkeit lexikalischer Vertreter in einem Paradigma. Wortverbindungen beziehen sich auf ihre systematische Fixiertheit im ganzen Syntagma. Usuelle Wortverbindungen bilden das zu Grunde liegende Schema ab, mit der entsprechenden Funktion, die es in der Kommunikation realisiert (vgl. „semantisch-pragmatische Einheiten“ bei Siepmann 2002, 250–252). Um zu den konstitutiven Bildungsmustern zu gelangen, muss man iterativ vorgehen: Ermittelte rekurrente Syntagmen werden wiederum daraufhin analysiert, in welchen Umgebungsmustern sie eingebettet sind. Auch die Festigkeit des ursprünglichen Ausgangsworts kann sich auf diese Weise relativieren. Folgende zwei Beispiele sollen verdeut-

¹¹ Schemann setzt in diesem Sinne tatsächlich *Grund*_{Singular}, *Gründe*, *Gründen* als getrennte Lemmata an. (1993, 287–289).

lichen, welche verschiedenen Ebenen von Kollokations- und Wortverbindungsmustern bei inhärenter Funktionsbedeutung auf diese Weise rekonstruierbar sind:

(3)

Kookkurrenzrelation	Grund _{Singular} – sehen _{Lemma}
rekurrentes Syntagma:	<i>keinen Grund sehen</i>
Wortverbindungsmuster:	<i>kein X sehen</i>
X-Füller:	<i>Grund, Anlass, Sinn, Ausweg, Handlungsbedarf, Motiv</i>
X ist Platzhalter für:	sprachliche Einheiten, die auf die Motivation eines Menschen referieren, etwas zu tun
Funktionsbedeutung:	Zurückweisung eines von anderen Sprechern eigentlich als sinnvoll angesehenen Sachverhalts

(4)

Kookkurrenzrelation:	wissen _{Lemma} – warum _{Konnektor}
rekurrentes Syntagma:	<i>sehr wohl wissen, warum</i>
Wortverbindungsmuster:	<i>sehr wohl X</i>
X-Füller:	<i>bewusst sein; in der Lage sein; sich ... vorstellen können; zu unterscheiden wissen; bemerkt haben</i>
X ist Platzhalter für:	sprachliche Einheiten, die auf kognitive Fähigkeiten bzw. mentale Zustände referieren
Funktionsbedeutung:	Bestätigung, dass ein Sprecher Gründe, Umstände und Bedingungen eines Sachverhaltes kennt oder dass er sich der Schwierigkeit dieses Sachverhalts und möglicher Kritik durchaus bewusst ist

Beispiel (4) zeigt, dass es sich bei den lexikalischen Füllern nicht unbedingt um Vertreter einer traditionellen Wortklasse, z.B. ein und derselben Wortart handeln muss. Vielmehr können auch sehr unterschiedliche Versprachlichungen ein einheitliches Schema mit einer entsprechenden Funktionsbedeutung konstituieren. Erstrebenswert wäre sicherlich ein elektronisches ‚Wörterbuch usueller Bildungsmuster‘, das mit Hilfe dieser Methodik einen Zugang zu abstrakteren Konstitutionsmustern der Sprache ermöglichen würde, einen Zugang zu konzeptuellen Ordnungssystemen auch quer zu klassischen Wortarten und z.B. systematisiert nach funktionaler Relevanz der Kollokationen bzw. Wortverbindungen in der Kommunikation:

(5)

nicht einsehen, warum/wieso
kein X sehen
Ich wüsste nicht, warum
 [Zurückweisung]

warum um Himmels willen?
aus [Adjektiv_{Neg-Qual}]Gründen
aus welchen Gründen auch immer
weiß der Teufel, warum
schleierhaft, warum
 [Unverständnis]

*unter fadenscheinigen Gründen
wird man sich fragen lassen müssen, warum
wird zu erklären haben, warum*
[Kritik]

Solche korpusgesteuert erarbeiteten Ergebnisse lassen sich nicht ohne Weiteres in lexikografische Ergebnisse umsetzen. Man stößt ständig auf neue und unerwartete Zusammenhänge und Vernetzungen, die es zunächst zu interpretieren gilt, und das möglichst ohne zu stark einschränkende Vorannahmen. Auch eine konkrete Nutzungssituation könnte eher kontraproduktiv sein: Man läuft Gefahr, nicht nutzungsrelevante Beobachtungen vorschnell auszublenden und blind zu werden für die ungewöhnlichen, nicht erwarteten bzw. möglicherweise nicht ins Raster passenden usuellen Phänomene. Dies scheint alles wenig kompatibel zu sein mit den Erfordernissen einer an konkreten Bedürfnissen ausgerichteten Lexikografie oder auch Fremdsprachendidaktik. Daher handelt es sich hier eher um Basisempirie hart an den Korpusdaten, um Grundlagenforschung also. Allerdings wird auch in Bezug auf Lexikografie und Fremdsprachendidaktik des Deutschen genau diese Grundlagenforschung in zunehmendem Maße eingefordert. Dass wir noch ziemlich am Anfang stehen, sollte uns jedoch nicht abhalten, schon jetzt korpuslinguistische Erkenntnisse für die praktische lexikografische Arbeit zu nutzen, wo immer es möglich ist.

Zu folgenden Aspekten der Kookkurrenzperspektive lässt sich Gewinn für moderne Kollokationswörterbücher (über das Selektionskriterium hinausgehend) erwarten:

1. *Lesarten des Bezugswortes.* Wenn die Einträge zu Basen in einem Kollokationswörterbuch nach Bedeutungen unterteilt werden sollen, liefern die Kookkurrenzen das entscheidende Kriterium: So müsste beim Stichwort *Grund*_{Lemma} als primäre Bedeutungen ‚Motiv/Ursache‘ und ‚ausgedrückte Sprachhandlungen‘ angegeben werden, da diese Verwendungsweisen die häufigsten sind. Die in vielen Wörterbüchern aufgeführten Bedeutungen wie 1. ‚Erdboden‘ 2. ‚kl. Tal‘ 3. ‚Boden eines Gewässers‘ spielen in den Kookkurrenzen nur eine sekundäre Rolle.¹²
2. *Mehrgliedrigkeit.* Die Perspektive der Zweigliedrigkeit von Kollokationen, ausgehend von der Basis, ist vor allem für wirkliches Wortschatzlernen (Lernen von zusammengehörenden Vokabeln) oder bei der Suche nach den adäquaten lexikalischen Partnern (als Reservoir stilistischer Varianten) sinnvoll. Bei zweigliedrigen Kollokationsangaben sollten nur solche vermerkt werden, die auch ohne syntagmatischen Kontext autonom verstehbar sind (siehe expressive Adjektive zu *Unterschied* in 4.; bewertende Adjektive zu *Grund*_{Lemma} in Abbildung 2). Auf komplexeren Aneignungs- bzw. Kompetenzstufen sind differenzierte Informationen zu den usualisierten syntagmatischen Einbettungen auch in produktionsorientierten Kollokationswörterbüchern m.E. unabdingbar (siehe *aus* [Adjektiv] *Gründen*). Hier kann die Kookkurrenzanalyse auch für die Ermittlung weiterer signifikanter Umgebungspartner der Wortverbindungen genutzt werden, die z. B. pragmatische Informationen liefern: z. B. *aus (durchaus, absolut) verständlichen Gründen*; (*aber, denn, doch*) *im Grunde*.

¹² Ausnahmen bilden hier LWDaF und PWDaF, die für *Grund* zwei Lemmata ansetzen (1. ‚Anlass/‘Motiv‘, 2. ‚Erdboden‘ usw.) und HWDaF, das die Lesart ‚Anlass/Motiv‘ an erster Stelle anordnet.

3. *Wortformenkollokationen*. Grundformenbasierte Kollokationsangaben dürften nur dann angesetzt werden, wenn sich die Fixiertheit auf das ganze Flexionsparadigma bezieht. Kollokationsrestriktionen wie **ein politischer/persönlicher*!... *Grund*_{Singular} sollten entsprechend vermerkt werden.
4. *Varianz und Festigkeit*. Varianzangaben sind als obligatorisch (<immer>) oder skalar fakultativ (z. B. <häufig>, <auch>) zu markieren, z. B.: *ungeklärt* – *Grund*_{Plural} –> <häufig> *aus ungeklärten Gründen* –> <auch>: aus [Temporaladverb z. B. *bisher, bislang, noch*] *ungeklärten Gründen* vs. **aus geklärten Gründen* –> *aus* [<immer> Negationsträger z. B. *nicht/nie*] *geklärten Gründen* –> <häufig> aus [Temporaladverb *bisher, bislang, noch*] *nicht/nie* <auch> [Modifikator, z. B. *ganz, vollständig*] *geklärten Gründen*.
5. *Relativ feste Wortverbindungen* (siehe Fall a. ‚Exklusivität‘ oder Fall b. ‚Dominanz‘) sollten auch in einem Kollokationswörterbuch Platz finden, gesondert behandelt und mit entsprechenden konkurrenzbasierten Informationen zum typischen Gebrauch angereichert werden, z. B. durch automatisch ermittelte externe Ergänzungen wie für *Grund und Boden* (*städtischer, öffentlicher G. und B., Erwerb/Ausverkauf* von G. und B., in G. und B. *blamieren/ schämen/ verdammen/ kritisieren; fahren/ laufen/ spielen/ reden/ stampfen/ wirtschaften/ bomben; förmlich/ regelrecht*). Wenn sich die Festigkeit einer Verbindung auf einer höheren Abstraktionsstufe relativiert, wären Hinweise auf das eigentliche Schema sehr interessant: *aus welchen Gründen auch immer* -> <gehört zu> *X auch immer* -> *X= aus welchen Gründen/Motiven (auch immer); in welcher Form/ Art und Weise (auch immer); für welchen Zweck (auch immer); warum/wann/wieso (auch immer)*.
6. *Onomasiologische Ordnungsprinzipien*. Eine Systematik wie in Abbildung 2 kann z. B. für onomasiologische Anordnungen von Kollokatoren im Wörterbuch hilfreich sein.

6. Nachbemerkung

Trotz der augenfälligen empirischen Revolution durch die Korpuslinguistik begegnet man immer wieder Skepsis gegenüber dem Nutzen automatischer Verfahren für lexikografische und fremdsprachendidaktische Zwecke (stellvertretend für diesen Diskurs vgl. Steinbügl 2005). Der Grund liegt jedoch m. E. in einer falschen Erwartungshaltung: Automatische Analysemethoden lösen nicht unsere lexikografischen Probleme. Mit ihnen werden Strukturen berechnet, unbestechlich, emotionsfrei und ohne didaktischen Hintergrund. Die Interpretation bleibt nach wie vor Angelegenheit des Menschen. Die neuen empirischen Möglichkeiten machen eine kluge Aufarbeitung der Befunde oftmals nicht einfacher, sondern komplexer und schwieriger. Kein Forschungsvorhaben und kein Wörterbuch werden allein in der Lage sein, auf dieser Basis das Wortverbindungsinventar einer Sprache umfassend linguistisch zu beschreiben und entsprechend zu dokumentieren. Es kann eigentlich nur einen einzigen Weg geben, diesen mächtigen Verfahren wirklich Neues für Forschung und Wörterbücher abzugewinnen: Das ist die Befähigung, mit den uns zur Verfügung stehenden Ressourcen und Analysemethoden reflektiert umzugehen. Die Schlagwörter heißen: Methodenkompetenz und Vernetzung. Ein wichtiges Ziel muss sein, diejenigen, die professionell mit der deutschen Sprache umgehen, dazu zu befähigen, Korpusbefunde ertragreich zu

‚lesen‘ und sie in ihrem jeweiligen Erkenntnisinteresse entsprechend auszuwerten. Die Verbindung von theoretischer Reflexion und anwendungsbezogenem Denken ist dafür genauso unabdingbar wie die Zusammenarbeit muttersprachlicher und nichtmuttersprachlicher Experten. Als hoffnungsvollen Ausblick für das Deutsche sollen Hausmanns Worte auf der IDS-Tagung 2003 diesen Beitrag beschließen:

„Die Zukunft gehört den Corpora. [...]. Das Deutsche auf der Basis der Corpora für den Rest der Welt brauchbar zu beschreiben, brauchbare Grammatiken und Wörterbücher (vor allem zweisprachige) des Deutschen für die Sprecher anderer Sprachen bereitzustellen – damit tut man etwas für das Überleben der Sprache in der Welt von Morgen.“ (2004, 322)

7. Bibliografie

7.1 Wörterbücher

- Agricola = Erhard Agricola (Hg.): Wörter und Wendungen. Wörterbuch zum deutschen Sprachgebrauch. Unter Mitwirkung v. Herbert Görner und Ruth Küfner. Leipzig 1962.
- DDUW = Duden Deutsches Universalwörterbuch. 6., überarb. u. erw. Aufl. Hrsg. v. der Dudenredaktion. Mannheim, Leipzig, Wien, Zürich 2007.
- DGWDS = Duden Das große Wörterbuch der deutschen Sprache in zehn Bänden. 3., völlig neu bearb. u. erw. Aufl. Hrsg. v. Wiss. Rat der Dudenredaktion. Band 4: Gele – Impr. Mannheim, Leipzig, Wien, Zürich 1999.
- Duden 2 = Duden Das Stilwörterbuch. 8., völlig neu bearb. Aufl. Hrsg. v. der Dudenredaktion. Mannheim, Leipzig, Wien, Zürich 2001.
- Duden 11 = Duden Redewendungen. Wörterbuch der deutschen Idiomatik. 2., neu bearb. u. aktual. Aufl. Hrsg. v. der Dudenredaktion. Mannheim, Leipzig, Wien, Zürich 2002.
- HwDaF = Hueber Wörterbuch Deutsch als Fremdsprache. Das einsprachige Wörterbuch für Kurse der Grund- und Mittelstufe. Erarbeitet von der Dudenredaktion und der Redaktion Deutsch als Fremdsprache des Hueber Verlags. 2. Auflage. Ismaning, Mannheim 2007.
- Kempcke = Günter Kempcke: De Gruyter Wörterbuch Deutsch als Fremdsprache. Berlin, New York 2000.
- LGWDaF = Langenscheidt Großwörterbuch Deutsch als Fremdsprache. Das einsprachige Wörterbuch für alle, die Deutsch lernen. Neubearbeitung. Hrsg. v. Dieter Götz, Günther Haensch und Hans Wellmann. In Zusammenarbeit mit der Langenscheidt-Redaktion. Berlin, München, Wien, Zürich, New York 2008.
- NODE = The New Oxford Dictionary of English. Hrsg. v. Judy Persall. Oxford University Press. Oxford 1998.
- OCD = Oxford Collocations Dictionary for Students of English. Oxford University Press. Oxford 2002.
- PGWDaF = PONS Grosswörterbuch Deutsch als Fremdsprache. 1. Auflage. Stuttgart 2006.
- Schemann = Hans Schemann: PONS Deutsche Idiomatik: Die deutschen Redewendungen im Kontext. Stuttgart, Dresden 1993.
- WDG = Wörterbuch der deutschen Gegenwartssprache. Hrsg. v. Ruth Klappenbach und Wolfgang Steinitz. 5. Aufl., 3. Band: glauben – Lyzeum. Berlin 1981.
- WDW = Wahrig Deutsches Wörterbuch. Hrsg. v. Renate Wahrig-Burfeind. 8., vollständig neu bearb. und aktualisierte Aufl. Gütersloh, München 2006.

7.2 Internetadressen (Zuletzt besucht am 29.2.2008)

- Belica 1995 = Cyril Belica: Statistische Kollokationsanalyse und Clustering. Korpuslinguistische Analyse-methode. Institut für Deutsche Sprache, Mannheim. (online abrufbar unter CII).
- Belica 2001–2007 = Cyril Belica: Kookkurrenzdatenbank CCDB – V3.2
Eine korpuslinguistische Denk- und Experimentierplattform für die Erforschung und theoretische Begründung von systemisch-strukturellen Eigenschaften von Kohäsionsrelationen zwischen den Konstituenten des Sprachgebrauchs. Institut für Deutsche Sprache, Mannheim. <http://corpora.ids-mannheim.de/ccdb/>
- CII = COSMAS II: <http://www.ids-mannheim.de/cosmas2/>
- DWDS = Das Digitale Wörterbuch der deutschen Sprache des 20. Jh.: <http://www.dwds.de/>
- KLa = Das Deutsche Referenzkorpus DeReKo:
<http://www.ids-mannheim.de/kl/projekte/korpora/>
- KLb = Methoden der Korpusanalyse und –erschließung:
<http://www.ids-mannheim.de/kl/projekte/methoden/>
- OWID = Online-Wortschatz-Informationssystem Deutsch des Instituts für Deutsche Sprache, Mannheim (OWID) am IDS: <http://www.owid.de>
- UWV = Usuelle Wortverbindungen: <http://www.ids-mannheim.de/ll/uwv/>
- Wortschatz = Wortschatzportal der Universität Leipzig: http://wortschatz.informatik.uni-leipzig.de/html/inhalt_next.html

7.3 Sekundärliteratur

- Agricola 1982 = Erhard Agricola: Ein Modellwörterbuch lexikalisch-semantischer Strukturen. In: Erhard Agricola, Joachim Schildt, Dieter Viehweger (Hg.): Wortschatzforschung heute. Aktuelle Probleme der Lexikologie und Lexikographie. Leipzig 1982, 9–22.
- Bahns 1996 = Jens Bahns: Kollokationen als lexikographisches Problem. Eine Analyse allgemeiner und spezieller Lernerwörterbücher des Englischen. Tübingen 1996 (Lexicographica. Series Maior 74).
- Belica/Steyer = Cyril Belica/Kathrin Steyer i. Dr.: Korpusanalytische Zugänge zu sprachlichem Usus. In: AUC (Acta Universitatis Carolinae). Germanistica Pragensia XX. Prag. (Vorabdruck unter <http://www.ids-mannheim.de/ll/uwv/publikationen.html>).
- Blumenthal 2006 = Peter Blumenthal: Wortprofil im Französischen. Tübingen 2006. (Beihefte zur Zeitschrift für Romanische Philologie 332).
- Brunner/Steyer 2007 = Annelen Brunner/Kathrin Steyer: Corpus-driven study of multi-word expressions based on collocations from a very large corpus. In: Proceedings of the 4th Corpus Linguistics conference, Birmingham. http://corpus.bham.ac.uk/corplingproceedings07/paper/182_Paper.pdf
- Bubenhof 2008 = Noah Bubenhof: „Es liegt in der Natur der Sache ...“ Korpuslinguistische Untersuchungen zu Kollokationen in Argumentationsfiguren. In: Carmen Mellado Blanco (Hg.): Beiträge zur Phraseologie aus textueller Sicht. Hamburg 2008, 53–72.
- Burger 2007 = Harald Burger: Phraseologie. Eine Einführung am Beispiel des Deutschen. 3. neu bearbeitete Auflage. Berlin 2007 (Grundlagen der Germanistik 36).
- Burger/Dobrovol'skij/Kühn/Norrick (Hg.) 2007a = Harald Burger/Dmitrij Dobrovol'skij/Peter Kühn/Neal R. Norrick: Phraseologie/Phraseology. Ein internationales Handbuch zeitgenössischer Forschung/An International Handbook of Contemporary Research. 2 Halbbände. Berlin, New York 2007 (HSK 28.1/2).
- Burger/Dobrovol'skij/Kühn/Norrick 2007b = Harald Burger/Dmitrij Dobrovol'skij/Peter Kühn/Neal R. Norrick: Phraseologie: Objektbereich, Terminologie und Forschungsschwerpunkte. In: Burger/Dobrovol'skij/Kühn/Norrick 2007a, 1–10.
- Busch/Stenschke 2007 = Albert Busch/Oliver Stenschke: Germanistische Linguistik. Eine Einführung. Tübingen 2007 (bachelor-wissen).
- Butina-Koller 2005 = Ekaterina Butina-Koller: Kollokationen im zweisprachigen Wörterbuch. Zur Behandlung lexikalischer Kollokationen in allgemeinsprachlichen Wörterbüchern des Sprachen-paares Französisch/Russisch. Tübingen 2005 (Lexicographica. Series Maior 124).

- Caro Cedillo 2004 = Ana Caro Cedillo: Fachsprachliche Kollokationen. Ein übersetzungsorientiertes Datenbankmodell Deutsch-Spanisch. Tübingen 2004 (Forum für Fachsprachen-Forschung 63).
- Coseriu 1967 = Eugenio Coseriu: „Lexikalische Solidaritäten“. In: *Poetica* 1, 293–303.
- Cruse 1986 = David Cruse: *Lexical semantics*. Cambridge 1986.
- Deppermann 2006 = Arnulf Deppermann: *Construction Grammar* – Eine Grammatik für die Interaktion? In: Arnulf Deppermann, Reinhard Fiehler, Thomas Spranz-Fogasy (Hg.): *Grammatik und Interaktion. Untersuchungen zum Zusammenhang zwischen grammatischen Strukturen und Gesprächsprozessen*. Radolfzell 2006. 43–65.
- Duffner 2006 = Rolf Duffner: Satzadverbien – korpusbasiert. In: Näf/Duffner 2006, 133–155.
- Duffner/Näf 2006 = Rolf Duffner/Anton Näf: Digitale Textdatenbanken im Vergleich. In: Näf/Duffner 2006, 7–22.
- Eichinger 2004 = Ludwig M. Eichinger: Von Köpfen, Nägeln und anderen guten Bekannten. In: Steyer 2004a, 1–15.
- Engelberg/Lemnitzer 2001 = Stefan Engelberg/Lothar Lemnitzer: *Lexikographie und Wörterbuchbenutzung*. Tübingen 2001 (Stauffenburg Einführungen 14).
- Feilke 1998 = Helmut Feilke: Idiomatic Prägnanz. In: Irmhild Barz, Günther Öhlschläger (Hg.): *Zwischen Grammatik und Lexikon*. Tübingen 1998, 69–80 (*Linguistische Arbeiten* 390).
- Feilke 2004 = Helmuth Feilke: Kontext – Zeichen – Kompetenz. Wortverbindungen unter sprachtheoretischem Aspekt. In: Steyer 2004a, 41–64.
- Fellbaum 2007 = Christiane Fellbaum (Hg.): *Idioms and Collocations. Corpus-based Linguistic and Lexicographic Studies*. London 2007 (*Research in Corpus and Discourse*).
- Firth 1957 = John Rupert Firth: *A Synopsis of Linguistic Theory 1930–1955*. In: *Studies in Linguistic Analysis*. Philological Society. Oxford 1957. Reprinted in Palmer, F. (ed. 1968) *Selected Papers of J. R. Firth*. Longman. Harlow.
- Fischer/Stefanowitsch 2006a = Kerstin Fischer/Anatol Stefanowitsch (Hg.): *Konstruktionsgrammatik. Von der Anwendung zur Theorie*. Tübingen 2006 (= *Stauffenburg Linguistik* 40).
- Fischer, Kerstin/Stefanowitsch, Anatol (2006b): *Konstruktionsgrammatik. Ein Überblick*, In: Fischer/Stefanowitsch 2006a, 3–17.
- Gladysz 2003 = Marek Gladysz: *Lexikalische Kollokationen in deutsch-polnischer Konfrontation*. Frankfurt/M., Berlin, Bern, Bruxelles, New York, Oxford, Wien 2003 (*Danziger Beiträge zur Germanistik* 11).
- Haß 2005 = Ulrike Haß (Hg.): *Grundfragen der elektronischen Lexikographie. elexiko – das Online-Informationssystem zum deutschen Wortschatz*. Berlin, New York 2005 (*Schriften des Instituts für Deutsche Sprache* 12).
- Hausmann 1984 = Franz Josef Hausmann: „Wortschatzlernen ist Kollokationslernen. Zum Lehren und Lernen französischer Wortverbindungen.“ In: *Praxis des neusprachlichen Unterrichts*, Jg. 31, S. 395–406.
- Hausmann 1985 = Franz Josef Hausmann: Kollokationen im deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels. In: Henning Bergenholtz, Joachim Mugdan (Hg.): *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch 28.–30.6.1984*. Tübingen 1985, 118–129 (*Lexicographica. Series Maior* 3).
- Hausmann 1989 = Franz Josef Hausmann: *Le dictionnaire de collocations*. In: Hausmann/Reichmann/Wiegand/Zgusta 1989, 1010–1019.
- Hausmann 1997 = Franz Josef Hausmann: *Semiotaxis und Wörterbuch*. In: Klaus-Peter Konerding, Andrea Lehr (Hg.): *Linguistische Theorie und lexikographische Praxis. Symposiumsvorträge*. Heidelberg 1996, 171–179 (*Lexicographica. Series Maior* 82).
- Hausmann 2004 = Franz Josef Hausmann: Was sind eigentlich Kollokationen? In: Steyer 2004a, 309–334.
- Hausmann/Reichmann/Wiegand/Zgusta 1989 = Franz Josef Hausmann/Oskar Reichmann/Herbert Ernst Wiegand/Ladislav Zgusta (Hg.): *Wörterbücher. Ein internationales Handbuch zur Lexikographie*. Berlin, New York 1989 (HSK 5.1).
- Heid 2005 = Ulrich Heid: Corpusbasierte Gewinnung von Daten zur Interaktion von Lexik und Grammatik: Kollokation – Distribution – Valenz. In: Friedrich Lenz, Stefan J. Schierholz (Hg.): *Corpuslinguistik in Lexik und Grammatik*. Tübingen 2005, 97–122 (*Stauffenburg Linguistik* 37).
- Herbst 1998 = Thomas Herbst: *Langenscheidts Großwörterbuch Deutsch als Fremdsprache und die britische Lernerlexikographie*. In: Wiegand 1998, 20–33.

- Heringer 1999 = Hans Jürgen Heringer: Das höchste der Gefühle. Empirische Studien zur distributiven Semantik. Tübingen 1999 (Stauffenburg Linguistik 15).
- Holderbaum/Kornelius 2001 = Anja Holderbaum/Joachim Kornelius: Kollokationen als Problemgrößen der Sprachmittlung. In: Andrea Lehr, Matthias Kammerer, Klaus-Peter Konerding, Angelika Storrer, Caja Thimm, Werner Wolski (Hg.): Sprache im Alltag: Beiträge zu neuen Perspektiven in der Linguistik. Herbert Ernst Wiegand zum 65. Geburtstag gewidmet. Tübingen 2001, 533–545.
- Hollós 2004 = Zita Hollós: Lernerlexikographie: syntagmatisch. Konzeption für ein deutsch-ungarisches Lernerwörterbuch. Tübingen 2004 (Lexicographica Series Maior 116).
- Irsula Peña 1994 = Jesús Irsula Peña: Substantiv-Verb-Kollokationen. Kontrastive Untersuchungen Deutsch–Spanisch. Frankfurt/M., Berlin, Bern, New York, Paris, Wien 1994 (Hispano-Americana 3).
- Kämper/Eichinger 2007 = Heidrun Kämper, Ludwig M. Eichinger (Hg.): Sprach-Perspektiven. Germanistische Linguistik und das Institut für Deutsche Sprache. Tübingen 2007, 493–509 (Studien zur deutschen Sprache 40).
- Keibel/Belica 2007 = Holger Keibel/Cyril Belica: CCDB: A Corpus-Linguistic Research and Development Workbench. In: Proceedings of the 4th Corpus Linguistics conference, Birmingham. http://corpus.bham.ac.uk/corplingproceedings07/paper/134_Paper.pdf.
- Klosa 2008 = Annette Klosa (Hg.) (2008): Lexikografische Portale im Internet. Mannheim 2008.
- Köster/Neubauer 2002 = Lutz Köster/Fritz Neubauer: Kollokationen und Kompetenzbeispiele im DeGruyter Wörterbuch Deutsch als Fremdsprache. In: Wiegand 2002, 283–310.
- Konerding 1998 = Klaus-Peter Konerding: Die semantischen Angaben in Langenscheidts Grosswörterbuch Deutsch als Fremdsprache. In: Wiegand 1998, 107–143.
- Kornelius 1995 = Joachim Kornelius: Vom Printwörterbuch zum elektronischen Kollokationswörterbuch. Theoretische, methodische und praktische Überlegungen zur Erstellung eines Kollokationswörterbuchs. In: Fredric F. M. Dolezal, Alain Rey, Thorsten Roelcke, Herbert Ernst Wiegand, Werner Wolski, Ladislav Zgusta (Hg.): Lexicographica 11. Tübingen 1995, 153–199.
- Lehr 1996 = Andrea Lehr: Kollokationen und maschinenlesbare Korpora. Ein operationales Analysemodell zum Aufbau lexikalischer Netze. Tübingen 1996 (Reihe Germanistische Linguistik 168).
- Lehr 1998 = Andrea Lehr: Kollokationen in Langenscheidts Großwörterbuch Deutsch als Fremdsprache. In: Wiegand 1998, 256–281.
- Lemnitzer 1997 = Lothar Lemnitzer: Akquisition komplexer Lexeme aus Textkorpora. Tübingen 1997 (Reihe Germanistische Linguistik 180).
- Lemnitzer/Zinsmeister 2006 = Lothar Lemnitzer/Heike Zinsmeister: Korpuslinguistik. Eine Einführung. Tübingen 2006 (Narr Studienbücher).
- Meibauer/Demske/Geilfuß-Wolfgang/Pafel/Ramers/Rothweiler/Steinbach 2007 = Jörg Meibauer/Ulrike Demske/Jochen Geilfuß-Wolfgang/Jürgen Pafel/Karl Heinz Ramers/Monika Rothweiler/Markus Steinbach: Einführung in die germanistische Linguistik. 2., aktualisierte Auflage. Stuttgart, Weimar 2007.
- Müller-Spitzer 2007 = Carolin Müller-Spitzer: Der lexikografische Prozess. Konzeption für die Modellierung der Datenbasis. Tübingen 2007 (Studien zur Deutschen Sprache 42).
- Näf/Duffner 2006 = Anton Näf/Rolf Duffner: Korpuslinguistik im Zeitalter der Textdatenbanken/Corpus linguistics in the era of text data banks (Linguistik online 28, 3/06).
- Perkuhn 2007 = Rainer Perkuhn: „Corpus-driven“: Systematische Auswertung automatisch ermittelter sprachlicher Muster. In: Kämper/Eichinger 2007, 465–491 (Studien zur deutschen Sprache 40).
- Porzig 1934 = Walter Porzig: „Wesenhafte Bedeutungsbeziehungen“. In: Beiträge zur Geschichte der deutschen Sprache und Literatur 58, 70–97.
- Reder 2006 = Anna Reder: Kollokationen in der Wortschatzarbeit. Wien 2006.
- Roos 1975 = Eckhard Roos: Kollokationsmöglichkeiten der Verben des Sehvermögens im Deutschen und Englischen. Bern, Frankfurt/M. 1975 (Forum Linguisticum 4).
- Sailer 2007 = Manfred Sailer: Corpus linguistic approaches with German corpora. In: Burger/Dobrovolskij/Kühn/Norrick (Hg.) 2007a, Halbband 2, 1060–1071.
- Schafroth 2003 = Elmar Schafroth: Kollokationen im GWDS. In: Wiegand 2003, 397–412.
- Scherfer 2002 = Peter Scherfer: Lexikalische Kollokationen. In: Ingo Kolboom, Thomas Kotschi, Edward Reichel (Hg.): Handbuch Französisch. Sprache – Literatur – Kultur – Gesellschaft. Für Studium, Lehre, Praxis. Berlin 2002, 230–237.

- Siepmann 2002 = Dirk Siepmann: „Eigenschaften und Formen lexikalischer Kollokationen: Wider ein zu enges Verständnis.“ In: Zeitschrift für französische Sprache und Literatur. Band 112. Heft 1, 240–263.
- Siepmann 2005 = Dirk Siepmann: „Collocation, Colligation and Encoding dictionaries. Part I.: Lexicological Aspects“. In: International Journal of Lexicography. Volume 18, Number 4. Oxford 2005, 409–443.
- Sinclair 1991 = John Sinclair: Corpus, Concordance, Collocation. Oxford 1991.
- Steinbügl 2005 = Birgit Steinbügl: Deutsch-englische Kollokationen. Erfassung in zweisprachigen Wörterbüchern und Grenzen der korpusbasierten Analyse. Tübingen 2005 (Lexicographica. Series Maior 126).
- Steyer 1998 = Kathrin Steyer: Kollokationen als zentrales Übersetzungsproblem – Vorschläge für eine Kollokationsdatenbank Deutsch-Französisch/Französisch-Deutsch auf der Basis paralleler und vergleichbarer Korpora. In: Daniel Bresson (Hg.): Lexikologie und Lexikographie Deutsch-Französisch. Aix-en-Provence 1998, 95–113 (Cahiers d'Études Germaniques 35).
- Steyer 2000 = Kathrin Steyer: „Usuelle Wortverbindungen des Deutschen. Linguistisches Konzept und lexikografische Möglichkeiten.“ In: Deutsche Sprache. Jahrgang 28, 2/2000, 101–125.
- Steyer 2003 = Kathrin Steyer: Korpus, Statistik, Kookkurrenz. Lässt sich Idiomatisches „berechnen“? In: Harald Burger, Annelies Häcki Buhofer, Gertrud Gréciano (Hg.): Flut von Texten – Vielfalt der Kulturen. Ascona 2001 zur Methodologie und Kulturspezifität der Phraseologie. Baltmannsweiler 2003, 33–46 (Phraseologie und Parömiologie 14).
- Steyer 2004a = Kathrin Steyer (Hg.): Wortverbindungen – mehr oder weniger fest. Berlin, New York 2004 (Jahrbuch des Instituts für Deutsche Sprache 2003).
- Steyer 2004b = Kathrin Steyer: Kookkurrenz. Korpusmethodik, linguistisches Modell, lexikografische Perspektiven. In: Steyer 2004a, 87–116.
- Steyer i. Dr. = Kathrin Steyer: Zwischen theoretischer Modellierung und praxisnaher Anwendung. Zur korpusgesteuerten Beschreibung usueller Wortverbindungen. In: Carmen Mellado Blanco (Hg.): Theorie und Praxis der idiomatischen Wörterbücher [Reihe Lexicographica Series Maior]. Tübingen 2009.
- Steyer/Lauer 2007 = Kathrin Steyer/Meike Lauer: „Corpus-Driven“: Linguistische Interpretation von Kookkurrenzbeziehungen. In: Kämper/Eichinger 2007, 493–509.
- Steyer/Lauer/Brunner 2008 = Kathrin Steyer/Meike Lauer/Annellen Brunner: Usuelle Wortverbindungen in OWID. Lexikografische Aspekte unter besonderer Berücksichtigung der Pragmatik. In: Kloss 2008.
- Stubbs 2001 = Michael Stubbs: Words and Phrases. Corpus Studies of Lexical Semantics. Oxford 2001.
- Tognini-Bonelli 2001 = Elena Tognini-Bonelli: Corpus Linguistics at Work. Amsterdam, Philadelphia 2001 (Studies in Corpus Linguistics 6).
- Viehweger, Dieter et al. (1977): Probleme der semantischen Analyse. Berlin 1977. (studia grammatica XV).
- Viehweger 1989 = Dieter Viehweger: Probleme der Beschreibung semantischer Vereinbarkeitsrelationen im allgemeinen einsprachigen Wörterbuch. In: Hausmann/Reichmann/Wiegand/Zgusta 1989, 888–893.
- Wiegand 1990 = Herbert Ernst Wiegand: Die deutsche Lexikographie der Gegenwart. In: Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand, Ladislav Zgusta: Wörterbücher. Ein internationales Handbuch zur Lexikographie. Berlin, New York 1990, 2100–2246 (HSK 5.2).
- Wiegand 1998 = Herbert Ernst Wiegand (Hg.): Perspektiven der pädagogischen Lexikographie des Deutschen. Untersuchungen anhand von „Langenscheidts Großwörterbuch Deutsch als Fremdsprache“. Tübingen 1998 (Lexicographica. Series Maior 86).
- Wiegand 2002 = Herbert Ernst Wiegand (Hg.): Perspektiven der pädagogischen Lexikographie des Deutschen II. Untersuchungen anhand des „de Gruyter Wörterbuchs Deutsch als Fremdsprache“. Tübingen 2002 (Lexicographica. Series Maior 110).
- Wiegand 2003 = Herbert Ernst Wiegand (Hg.): Untersuchungen zur kommerziellen Lexikographie der deutschen Gegenwartssprache I. „Duden. Das große Wörterbuch der deutschen Sprache in zehn Bänden“. Print- und CD-ROM-Version. Tübingen 2003 (Lexicographica. Series Maior 113).
- Wiegand 2005 = Herbert Ernst Wiegand (Hg.): Untersuchungen zur kommerziellen Lexikographie der deutschen Gegenwartssprache II. „Duden. Das große Wörterbuch der deutschen Sprache in

- zehn Bänden“. Print- und CD-ROM-Version. Band 2. Tübingen 2005 (Lexicographica. Series Maior 121).
- Zöfgen 2001 = Ekkehard Zöfgen: Lexikalische Zweierbeziehungen: ‚Vertraute Unbekannte‘ im mentalen Lexikon germanophoner Französischlerner. In: Karin Aguado, Claudia Riemer (Hg.): Wege und Ziele. Zur Theorie, Empirie und Praxis des Deutschen als Fremdsprache (und anderer Fremdsprachen). Festschrift für Gert Henrici zum 60. Geburtstag. Hohengehren 2001, 267–286 (Perspektiven Deutsch als Fremdsprache).

